JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **126**, 455–468 (1987)

The Beta-Algorithm

EFIM A. GALPERIN

Département de mathématiques-informatique, Université du Québec à Montréal, C.P. 8888, Succ. A, Montréal, Québec, H3C 3P8, Canada

Submitted by George Leitmann

Received March 10, 1986

A set-monotonic non-gradient algorithm is proposed for finding global minima of general non-convex mathematical programming problems. The algorithm is based on the Cubic Algorithm (E. A. Galperin, *J. Math. Anal. Appl.* **112** (1985), 635–640) equipped with a semi-certain distinction operator and the marginal comparison constant generator.

1. INTRODUCTION

We consider the problem

$$\min f(x), \qquad x \in \bar{X} \subset R^n, \tag{1.1}$$

of finding the value

$$p^\circ = \min_{x \in \bar{X}} f(x) \tag{1.2}$$

and the set

$$\bar{X}^\circ = \{x \mid f(x) = p^\circ, x \in \bar{X}\}, \tag{1.3}$$

where f is a continuous function and $\bar{X}$ is a compact robust set which may be non-convex and non-connected. For example, $\bar{X}$ may consist of a collection of closed tori, balls, and cubes that may intersect one another.

By definition, a set Y is called robust, if the closure of its interior coincides with its closure:

$$\text{cl int } Y = \text{cl } Y \tag{α}$$

(note that Y may be open, or closed, or neither open nor closed). A robust set $\bar{X}$ always has a non-empty interior, int $\bar{X} \neq \varnothing$. An open set X is always robust, its closure is denoted by $\bar{X}$ and its boundary by ∂X; clearly, int $\bar{X} =$

0022-247X/87 $3.00

$\text{int}\, X = X$. A closed set $\bar{Z} \subset R^n$ having empty interior, $\text{int}\, \bar{Z} = \varnothing$, is non-robust, examples: a point, an arc in R^2, a circle in R^3. Let $Y = \bar{X} \cup Z$, $Z \not\subset \bar{X}$, where $\bar{X} \subset R^n$ is a closed ball and $Z \subset R^n$, $n \geqslant 2$, an open arc. Then, $\text{int}\, Y = X$, $\text{cl int}\, Y = \bar{X}$, $\text{cl}\, Y = \bar{X} \cup \bar{Z}$, so that $\text{cl int}\, Y = \bar{X} \neq \bar{X} \cup \bar{Z} = \text{cl}\, Y$; thus, Y is non-robust although $\text{int}\, Y = X \neq \varnothing$.

In the sequel we shall need the following

LEMMA. *If $\bar{X} \subset R^n$ is robust and compact, $f\colon R^n \to R$ is continuous over $\bar{X}$, and $c = \text{const}$, then the set*

$$Y = \{x \mid f(x) < c,\ x \in \bar{X}\} \tag{β}$$

is either empty, or robust for every c.

Proof. Suppose that c is such that $Y \neq \varnothing$. Then we have from the definition of Y in (β): $Y \subseteq \bar{X}$. If $Y = \bar{X}$, then Y is robust and the statement follows. Consider such c that $Y \subset \bar{X}$. Since $\bar{X}$ is robust, the interior of Y is non-empty and is given (see (β)) by

$$\text{int}\, Y = \{x \mid f(x) < c, x \in X\} \neq \varnothing. \tag{γ}$$

For any set Y we have $Y \subseteq \text{cl}\, Y$. Since $\bar{X}$ is compact and f is continuous over $\bar{X}$ and because $Y \neq \varnothing$, from (β) it follows that the closure of Y is non-empty and is given by

$$\text{cl}\, Y = \{x \mid f(x) \leqslant c, x \in \bar{X}\}. \tag{δ}$$

On the other hand, by compactness and robustness of $\bar{X}$ and by the continuity of f over $\bar{X}$, we obtain from (γ)

$$\text{cl int}\, Y = \{x \mid f(x) \leqslant c, x \in \bar{X}\}. \tag{ε}$$

Comparing (δ) and (ε), we conclude (α), so that the set Y is robust. ∎

Since $\bar{X}$ is bounded, we can introduce a circumscribed (non-strictly) closed cube $\bar{C}$ such that

$$\bar{X} \subset \bar{C} \subset R^n. \tag{1.4}$$

Non-strict circumscription means that $\bar{X}$ and $\bar{C}$ may have no common boundary points. Let $c > 0$ be the length of the edge of $\bar{C}$, so that its volume is c^n.

Hypothesis. The cost function $f\colon R^n \to R$ is assumed to be a continuous single-valued computable procedure defined and Lipschitzian on $\bar{C}$, that is,

$$|f(x) - f(x')| \leqslant L\, \|x - x'\|; \qquad L = \text{const} > 0,\ x, x' \in \bar{C}. \tag{1.5}$$

If $\bar{X}=\bar{C}$, then under (1.5) the problem could be solved by the cubic algorithm the principal scheme of which, as described in [1], we reproduce briefly with minor generalization concerning the grid generator. Take a point $x_0 \in \bar{C}$ and an integer $N \geqslant 2$ and partition $\bar{C}$ into N^n equal subcubes $\bar{C}_i^1$ such that $C_i^1 \cap C_j^1 = \varnothing$ for $i \neq j$ and $\bigcup_{i=1}^{N^n} \bar{C}_i^1 = \bar{C}$. The volume of each $\bar{C}_i^1$ is c^n/N^n and its diameter (diagonal) is $c\sqrt{n}/N$:

$$d(C_i^1) = \sup_{x, x' \in \bar{C}_i^1} \|x - x'\| = \frac{c\sqrt{n}}{N}, \qquad \forall i. \tag{1.6}$$

It is clear that after partition the point $x_0 \in \bar{C}$ will be located in one (or more, if x_0 happens to be on a common boundary) of $\bar{C}_i^1$. In any case we *assign* x_0 to just one of $\bar{C}_i^1$, say, to $\bar{C}_{i_0}^1$ and we call x_0 the representative of $\bar{C}_{i_0}^1$. Apply parallel translation of $\bar{C}_{i_0}^1$ to make it coincide, turn by turn, with each $\bar{C}_i^1$, $i=1,..., N^n$, $i \neq i_0$; then $x_0 \in \bar{C}_{i_0}^1$ will define the representative $x_i^1 \in \bar{C}_i^1$ in each $\bar{C}_i^1$. This rule defines the "translated grid generator" and the collection of x_i^1, $i=1,..., N^n$ $(x_{i_0}^1 = x_0)$, yields the grid for any particular choice of $x_0 \in \bar{C}$.

The first comparison constant s_0 and the first deletion constant r_1 are

$$s_0 = f(x_0); \qquad r_1 = Ld(C_i^1) = \frac{Lc\sqrt{n}}{N}, \qquad \forall i. \tag{1.7}$$

Certain subcubes $\bar{C}_i^1 \subset \bar{C}$ will be further partitioned in the same way and subsequent deletion constants are

$$r_m = \frac{r_{m-1}}{N} = \frac{Lc\sqrt{n}}{N^m}, \qquad m = 2, 3,...,. \tag{1.8}$$

Iteration 1. Compute all $f(x_i^1)$, $i=1,..., N^n$. Delete all $\bar{C}_i^1$ for which

$$f(x_i^1) - s_0 > r_1. \tag{1.9}$$

The remaining subcubes correspond to the index set

$$I_1 = \{i \mid f(x_i^1) - s_0 \leqslant r_1,\ 1 \leqslant i \leqslant N^n\}. \tag{1.10}$$

The closure of those subcubes defines the set

$$\bar{K}_1 = \{x \mid x \in \bar{C}_i^1, i \in I_1\} \subseteq \bar{C}. \tag{1.11}$$

Compute

$$s_1 = \min_{i \in I_1} f(x_i^1), \tag{1.12}$$

which constitutes the extremal comparison constant generator. Clearly, $s_1 \leqslant s_0$.

Further iterations. Partition each $\bar{C}_i^1 \in \bar{K}_1$ in the same way as $\bar{C}$ and generate the new (finer) grid with the same translation rule. Repeat Iteration 1 replacing s_0, r_1 by s_1, r_2, then by s_2, r_3, etc. In this process we come to the two monotonic sequences

$$s_0 \geqslant s_1 \geqslant s_2 \geqslant \cdots \geqslant s_m \geqslant \cdots \tag{1.13}$$

$$\bar{C} \supseteq \bar{K}_1 \supseteq \bar{K}_2 \supseteq \cdots \supseteq \bar{K}_m \supseteq \cdots, \tag{1.14}$$

THE CONVERGENCE THEOREM [1].

$$\lim_{m \to \infty} s_m = s^{\circ} = \min_{x \in \bar{C}} f(x) \tag{1.15}$$

$$\lim_{m \to \infty} \bar{K}_m = \bigcap_{m=1}^{\infty} \bar{K}_m = \bar{K}^{\circ} = \{x \mid f(x) = s^{\circ}, x \in \bar{C}\}. \tag{1.16}$$

For the proof and more details see [1]. We need only mention two properties useful for the sequel:

(a) The translated grid generator guarantees that every grid point will remain in the iteration process until deleted with its corresponding subcube.

(b) The deletion operator (1.9) does not eliminate global minimizers (i.e., points of $\bar{K}^{\circ}$). In fact, we have (see [1])

$$\min_{x \in \bar{C}_i^1} f(x) > s_0, \qquad i \in I_0 - I_1,\ I_0 = \{1, 2, ..., N^n\} \tag{1.17}$$

for every subcube $\bar{C}_i^1$ deleted by (1.9), which means in terms of $\bar{K}_m$,

$$\inf_{x \in \bar{K}_{m-1} - \bar{K}_m} f(x) > s_{m-1}, \qquad m = 1, 2, ..., \bar{K}_0 = \bar{C}. \tag{1.18}$$

It is clear that the cubic algorithm does not solve the problem (1.1). Since $\bar{X} \subset \bar{C}$ (1.4) we have only

$$p^{\circ} \geqslant s^{\circ}. \tag{1.19}$$

If $p^{\circ} > s^{\circ}$, then obviously $\bar{X}^{\circ} \cap \bar{K}^{\circ} = \varnothing$; otherwise, by virtue of the First Alternative (see [2]) we have

$$\bar{X}^{\circ} = \bar{X} \cap \bar{K}^{\circ}. \tag{1.20}$$

In neither case does the cubic algorithm give the solution of (1.1). To solve the problem, we have to introduce certain new devices.

2. Application of the Cubic Algorithm to Quasi-cubic Sets

Definition 1. A set $\bar{S}$ is called quasi-cubic if it can be represented as a union of a finite number of closed cubes of equal volume:

$$\bar{S} = \bigcup_{i=1}^{M} \bar{C}_i, \qquad \text{vol}\, \bar{C}_i = c_i^n = \text{const},\ n = \dim \bar{C}_i. \tag{2.1}$$

The cubes $\bar{C}_i$ do not have to be connected nor uniformly oriented, and they may also intersect one another. A parallelepiped the ratios of any two edges of which are all rational is a quasi-cubic set; if there is at least one irrational ratio, it is not a quasi-cubic set. A simplex may be quasi-cubic or not. A ball is not a quasi-cubic set.

All sets $\bar{K}_m$ in (1.14) are quasi-cubic (not cubes!), thereby uniformly oriented (i.e., all constituent cubes have parallel edges). It is clear that, if in (1.1) we had $\bar{X} = \bar{K}_l$, the same algorithm would work although $\bar{K}_l$ might be not a cube. If $\bar{X}$ were a quasi-cubic set non-uniformly oriented but without intersections of constituent cubes, the same algorithm would also work. For a general quasi-cubic set the scheme may need minor modification with which it will also work.

This observation opens the way to the solution of the original problem (1.1) with a non-quasi-cubic set $\bar{X}$. Namely, if $\bar{X}$ can be approximated to any given precision by a quasi-cubic set, then the problem (1.1) can be solved through a version of the cubic algorithm. This, however, may be needlessly complex and we prefer to introduce some new devices to deal with optimization sets of general nature.

3. The Semi-certain Distinction Operator

To simplify the exposition, we use, wherever possible, the terms and notations already introduced.

Definition 2. Given $\bar{X}$, $\bar{C}_i$, a distinction operator is defined by the binary function

$$\varphi_i = \varphi(\bar{X}, \bar{C}_i) = \begin{cases} 0, & \text{if } \bar{X} \cap \bar{C}_i = \varnothing \text{ for sure} \\ 1, & \text{otherwise.} \end{cases} \tag{3.1}$$

Here the quantifier "for sure" has nothing to do with probability (not to be confused with "almost sure"); also, the term "otherwise" does not always mean $\bar{X} \cap \bar{C}_i \neq \varnothing$.

In some cases it is easy to construct distinction operators.

EXAMPLE 1. Suppose that $\bar{C} \subset R^n$ is a closed cube with the edge $c > 0$ and $\bar{X} \subset R^n$ is a finite union of closed balls:

$$\bar{X} = \{x \mid \|x - a_j\|^2 \leqslant b_j^2,\ b_j > 0,\ j = 1, \ldots, l\}. \tag{3.2}$$

For this case a distinction operator is given by the inequality

$$\|z - a_j\| > d_j, \qquad z \in \bar{C},\ j = 1, \ldots, l, \tag{3.3}$$

where

$$d_j \geqslant b_j + c\sqrt{n}, \tag{3.4}$$

with the understanding that $\varphi = 0$, if (3.3) holds for any $z \in \bar{C}$ fixed in advance, and $\varphi = 1$ otherwise. Here $\|\cdot\|$ is a Euclidean norm in R^n and $z \in \bar{C}$ is an arbitrary *fixed* point in $\bar{C}$. We see that a distinction operator is not unique and for the above case the one given by the equality in (3.4) may be considered as a "better" one.

To further clarify the sense of Definition 2, let us consider an example in one dimension.

EXAMPLE 2. Let $\bar{X} = [-2, +2]$, $\bar{C}_1 = [2, 3]$, $\bar{C}_2 = [2.5, 3.5]$, $\bar{C}_3 = [3, 4]$, and $\bar{C}_4 = [4, 5]$. Then we have $\varphi_1 = \varphi_2 = \varphi_3 = 1$ and $\varphi_4 = 0$, meaning that only $\bar{X} \cap \bar{C}_4 = \varnothing$ for sure. In fact, $\bar{X} \cap \bar{C}_1 \neq \varnothing$, $\bar{X} \cap \bar{C}_2 = \varnothing$, $\bar{X} \cap \bar{C}_3 = \varnothing$ but all three are *not* "for sure." Indeed, this example is a special case of (3.2), (3.3), (3.4) with $d \geqslant 2 + 1 = 3$, $n = 1$, $l = 1$, $a_1 = 0$, $b_1 = 2$, $c = 1$ for all four cubes. Our distinction operator is, thus,

$$\varphi_i = \begin{cases} 0, & \text{if } |z_i| > d \geqslant 3,\ z_i \in \bar{C}_i \qquad (3.5) \\ 1, & \text{otherwise.} \qquad (3.6) \end{cases}$$

Let $d = 3$ and take $z_1 = 3 \in \bar{C}_1$. For this z_1 we have (3.6), $\varphi_1 = 1$, since (3.5) is not satisfied.

If we took $d < 3$, then (3.5) would be satisfied, yielding $\varphi_1 = 0$, in contradiction with the fact that $\bar{X} \cap \bar{C}_1 = \{2\} \neq \varnothing$. This illustrates that inequalities (3.3), (3.4) are *unimprovable*.

Now, take $z_2 = 3 \in \bar{C}_2$, yielding, as before, $\varphi_2 = 1$ despite the fact that, had we taken $z_2' = 3.5 \in \bar{C}_2$, we would have (3.5) satisfied for this *particular* choice. The point here is that the statement "$\bar{X} \cap \bar{C}_i = \varnothing$ *for sure*" means that the empty intersection is to be established by the check based on one single point in $\bar{C}_i$ *arbitrarily fixed in advance*, which is essential for the operation of the Beta-Algorithm below. Of course, we cannot check all points $z \in \bar{C}_i$ that $z \notin \bar{X}$ to establish $\bar{X} \cap \bar{C}_i = \varnothing$ *for sure* for somewhat more general sets $\bar{X}$. That is why we need a distinction operator capable of separating $\bar{X}$ and $\bar{C}_i$ by the check of one single point $z_i \in \bar{C}_i$ *arbitrarily fixed* *in advance*. Naturally, such an operator cannot be fully certain and this is the sense of the word "otherwise" in (3.1) comprising all cases in which the non-intersection cannot be established by the check of a single point $z_i \in \bar{C}_i$. In the example, those are cases for $\bar{C}_1$, $\bar{C}_2$, $\bar{C}_3$ each of which has at least one point, e.g., $z = 3$, for all three cubes, violating (3.5).

The situation is different for $\bar{C}_4$. Here, whatever $z \in \bar{C}_4$, we always have $|z| > 3$ and, thus, $\varphi_4 = 0$.

Let us now take $d = 4 > 3$ in (3.5). In this case $\varphi_4 = 1$ (for $z = 4 \in \bar{C}_4$ the inequality $|z| > d$ is violated); however, for $\bar{C}_5 = [4.5, 5.5]$ we have $\varphi_5 = 0$, meaning that with such d the operator is still working, but it is poorer, having a greater uncertainty band.

From the above examples we can see that the distinction operator generates a compact set $\bar{Y}$ containing $\bar{X}$ and such that, given *any* fixed $z \in \bar{C}$, we have $\bar{X} \cap \bar{C} = \emptyset$, for sure, if $z \notin \bar{Y}$, denoted in (3.1) as $\varphi = 0$; otherwise, $\varphi = 1$, which corresponds to two possibilities: $z \in \bar{X}$, meaning $\bar{X} \cap \bar{C} \neq \emptyset$, and $z \in \bar{Y} - \bar{X}$, in which case nothing is known about the intersection. For a "good" distinction operator the sets $\bar{Y}$ and $\bar{X}$ are congruent such that for the cubic set $\bar{C} \subset R^n$ with the edge $c > 0$ the boundaries ∂Y and ∂X are $c\sqrt{n}$-equidistant surfaces, yielding the uniform uncertainty band $\mathscr{D} = \bar{Y} - \bar{X}$.

Let us consider variable sets $\bar{X}$, $\bar{C}$ parametrized by a parameter $m \in [0, \infty)$,

$$\bar{X}_m = \bar{X}(m), \qquad \bar{C}^m = \bar{C}(m), \tag{3.7}$$

where $\bar{X}(0) = \bar{X}$, $\bar{C}(0) = \bar{C}$, and all $\bar{C}^m$ are closed cubes with edges of the length

$$c_m = \frac{c}{N^m}, \qquad m = 0, 1, 2, \ldots. \tag{3.8}$$

In this case the uncertainty band $\mathscr{D}$ will depend on m: $\mathscr{D}_m = \mathscr{D}(m) = \bar{Y}_m - \bar{X}_m$.

Suppose $\bar{X}_{m+1} \subseteq \bar{X}_m$ $(m = 0, 1, \ldots)$ and that the intersection of all $\bar{X}_m$ is non-empty, which allows us to define the limit

$$\lim_{m \to \infty} \bar{X}_m = \bigcap_{m=0}^{\infty} \bar{X}_m = \bar{X}^{\circ} \neq \emptyset, \qquad \text{clearly } \bar{X}^{\circ} \subseteq \bar{X}. \tag{3.9}$$

Suppose also that $\lim_{m \to \infty} \bar{Y}_m$ is defined.

DEFINITION 3. A semi-certain distinction operator $\varphi(m) = \varphi(\bar{X}_m, \bar{C}^m)$ is called precise, if

$$\lim_{m \to \infty} \mathscr{D}_m = \emptyset, \tag{3.10}$$

or, equivalently,

$$\lim_{m \to \infty} \bar{Y}_m = \lim_{m \to \infty} \bar{X}_m = \bar{X}^{\circ}. \tag{3.11}$$

It means, of course, that the boundaries ∂Y_m and ∂X_m tend to coincide as $m \to \infty$. Clearly, the parameter m does not have to be discrete. One may choose a continuous parameter, for example, the length c of the edge of a variable cube $\bar{C}$.

Precise distinction operators are not unique. For instance, suppose that $\bar{X}_m = \bar{X} = \text{const}$ given by (3.2). If in (3.4) we take

$$d_j = b_j + Mc\sqrt{n}, \qquad M \geqslant 1, \tag{3.12}$$

then all operators (3.3) will be precise as $c \to 0$, however poor they may be for a big M and any fixed $c > 0$. On the contrary, the operator (3.3) with

$$d_j = d = \max_{1 \leqslant j \leqslant l} b_j + c\sqrt{n} \tag{3.13}$$

may be a "very good" one for any fixed $c > 0$ (e.g., if $\delta = \max b_j - \min b_j$ is small), but it is imprecise as $c \to 0$.

We have formulated the precision property with respect to the integer parameter $m = 0, 1, 2,..., m \to \infty$, since we shall need it in this form for the Beta-Algorithm. It is clear that the quality of "precision" describes the action of the distinction operator when $\bar{X}$ varies, but is always non-empty, and the comparison cube becomes smaller as its edge $c \to 0$.

One cannot propose a distinction operator good enough for all cases. In each case an operator should be constructed according to the problem under consideration.

4. The Beta-Algorithm

To facilitate comparisons, we employ the same notations as in [1] and in Section 1 when they concern the same objects. Starting with the same $\bar{X}$, $\bar{C}$ as in (1.4), and $N \geqslant 2$, we denote $\bar{B}_0 = \bar{C}$, $I_0 = \{1, 2,..., N^n\}$ and proceed as follows.

Iteration 1. Take $x_0 \in \bar{X}$ (not $x_0 \in \bar{C}$ as before) and partition $\bar{B}_0 = \bar{C}$ into N^n subcubes $\bar{C}_i^1$, such that $C_i^1 \cap C_j^1 = \varnothing$ for $i \neq j$ and $\bigcup \bar{C}_i^1 = \bar{C}$. Basing on x_0, generate the translated grid $x_i^1 \in \bar{C}_i^1$ as in the cubic algorithm (see Section 1) and compute all $f(x_i^1)$, $i \in I_0$.

Basing on x_i^1, apply a *precise* distinction operator $\varphi_i(c)$ and exclude from further considerations every $\bar{C}_i^1$ for which $\varphi_i = 0$, $i \in I_0$. The remaining subcubes correspond to the index set

$$J_0 = \{i \mid \varphi_i = 1,\ i \in I_0\}, \qquad \text{clearly } J_0 \neq \varnothing \tag{4.1}$$

(note that x_0 cannot be deleted by this operation). The closure of the subcubes with $i \in J_0$ defines the set

$$\bar{S}_0 = \{x \mid x \in \bar{C}_i^1,\ i \in J_0\} \subseteq \bar{B}_0, \qquad \bar{S}_0 \neq \varnothing. \tag{4.2}$$

Compute $p_0 = f(x_0)$ and delete all $\bar{C}_i^1 \in \bar{S}_0$ for which

$$f(x_i^1) - p_0 > r_1, \qquad i \in J_0,\ r_1 \text{ as in (1.7)}. \tag{4.3}$$

The remaining subcubes correspond to the index set

$$I_1^* = \{i \mid f(x_i^1) - p_0 \leqslant r_1,\ i \in J_0\}. \tag{4.4}$$

Here $I_1^* \subseteq I_1$, I_1 as given by (1.10), and $I_1^* \neq \varnothing$. The closure of the subcubes with $i \in I_1^*$ defines the set

$$\bar{B}_1 = \{x \mid x \in \bar{C}_i^1,\ i \in I_1^*\}, \qquad \bar{B}_1 \neq \varnothing. \tag{4.5}$$

Check the memebership $x_i^1 \in \bar{X}$ for each $i \in I_1^*$ and define the index set

$$J_1 = \{i \mid x_i^1 \in \bar{X}, i \in I_1^*\}. \tag{4.6}$$

Clearly, $J_1 \neq \varnothing$ since $x_0 \in \bar{X}$; also, $J_1 \subseteq I_1^* \subseteq J_0 \subseteq I_0$. Compute

$$p_1 = \min_{i \in J_1} f(x_i^1), \tag{4.7}$$

which constitutes the marginal comparison constant generator different from (1.12). Clearly, $p_1 \leqslant p_0$. Single out $x_1 \in \bar{X}$ such that $f(x_1) = p_1$ (it is made simultaneously with (4.7)).

Further iterations. Partition each $\bar{C}_i^1 \in \bar{B}_1$ in the same way as $\bar{B}_0$ and repeat Iteration 1, replacing x_0, p_0, r_1 by x_1, p_1, r_2, then by x_2, p_2, r_3, etc., with r_m given by (1.8). In this process we come to the two monotonic sequences

$$p_0 \geqslant p_1 \geqslant p_2 \geqslant \cdots \geqslant p_m \geqslant \cdots \tag{4.8}$$

$$\bar{C} = \bar{B}_0 \supseteq \bar{B}_1 \supseteq \bar{B}_2 \supseteq \cdots \supseteq B_m \supseteq \cdots. \tag{4.9}$$

THEOREM 1.

$$\lim_{m \to \infty} p_m = p^\circ = \min_{x \in \bar{X}} f(x) \tag{4.10}$$

$$\lim_{m \to \infty} \bar{B}_m = \bigcap_{m=0}^{\infty} \bar{B}_m = \bar{X}^\circ = \{x \mid f(x) = p^\circ, x \in \bar{X}\}. \tag{4.11}$$

Proof. To facilitate comparisons, we follow the same lines as in [1].

(a) *Existence and Nature of the Limit in R*

Since $\bar{X}$ is compact, there exists x^o such that $f(x^o)=p^o=\min_{x\in\bar{X}} f(x)$. Any distinction operator cannot exclude points of $\bar{X}$. By the choice $x_0\in\bar{X}$ and by construction of the translated grid generator and of the marginal comparison constant generator (4.6), (4.7), all basic points $x_0, x_1, ..., x_m \cdots$ belong to $\bar{X}$, and since $p_m=f(x_m)$, the sequence p_m in (4.8), monotonic and bounded from below by p^o, tends to a limit $\bar{p}\geqslant p^o$. There exists $\bar{x}\in\bar{X}$ such that $f(\bar{x})=\bar{p}$. Since $r_m=Lc\sqrt{n}/N^m\to 0$ as $m\to\infty$, then, if $\bar{p}>p^o$, the points $\bar{x}$ would be deleted at some stage, if two conditions hold:

(i) at least one global minimizer $x^o\in\bar{X}^o$ remains in the process indefinitely, that is, $\exists x^o$ such that $x^o\in\bar{B}_m$, $\forall m$;

(ii) the descent in (4.8) does not cease, i.e., the process cannot get stuck at some $p_l>p^o$ because of the non-appearance of new basic points x_m yielding $p_m=f(x_m)<p_l$.

In the case of $\bar{x}$ deleted, we would have $\lim_{m\to\infty} p_m<\bar{p}$, contrary to its definition: $\bar{p}=\lim_{m\to\infty} p_m$. Thus, to prove (4.10), $\bar{p}=p^o$, we have to show that conditions (i) and (ii) both hold.

(b) *Non-elimination of a Global Minimizer $x^o\in\bar{X}^o$*

No distinction operator can eliminate points of $\bar{X}$. The translated grid generator guarantees that every $x_i^m\in\bar{X}$ will remain in the iteration process until deleted with its corresponding subcube $\bar{C}_i^m$ by the deletion operator (4.3) (note that $x_i^m\notin\bar{X}$ may be deleted by the distinction operator). By (1.5), (1.6), (1.7) we have

$$\max_{x\in\bar{C}_i^1} f(x)-\min_{x\in\bar{C}_i^1} f(x)\leqslant L\,\|\arg\max_{x\in\bar{C}_i^1} f(x)-\arg\min_{x\in\bar{C}_i^1} f(x)\|$$
$$\leqslant L\max_{x,x'\in\bar{C}_i^1}\|x-x'\|=Ld(C_i^1)=r_1, \tag{4.12}$$

which implies

$$f(x_i^1)-p_0>r_1\geqslant\max_{x\in\bar{C}_i^1} f(x)-\min_{x\in\bar{C}_i^1} f(x)\geqslant f(x_i^1)-\min_{x\in\bar{C}_i^1} f(x) \tag{4.13}$$

so that

$$\min_{x\in\bar{C}_i^1} f(x)>p_0 \tag{4.14}$$

for every cube $\bar{C}_i^1$ deleted by (4.3). For those $\bar{C}_i^1$ containing portions of $\bar{X}$ the inequality (4.14) implies that

$$\min_{x\in\bar{C}_i^1\cap\bar{X}} f(x)>p_0, \tag{4.15}$$

the minimum existing by compactness of the intersection. The inequality

(4.15) demonstrates that deletion operator (4.3), acting on points $x_i^1 \in \bar{C}$ (not necessarily on points $x_i^1 \in \bar{X} \subset \bar{C}$!), cannot eliminate a global minimizer $x^\circ \in \bar{X}^\circ \subset \bar{X}$. This also proves that $\bar{X}^\circ \subset \bar{B}_m$ for all $m = 0, 1, \ldots$, whence $\bar{X}^\circ \subseteq \bigcap_{m=0}^{\infty} \bar{B}_m$.

(c) *Non-ceasing Descent in* (4.8)

Suppose, on the contrary, that descent does cease at some p_l, $l \geqslant 0$, because of the non-appearance of points $x_m \in \bar{X}$, $m > l$, such that (4.7) would yield $p_m = f(x_m) < p_l$. This implies that a portion $\bar{X}^* \subseteq \bar{X}$ for which $f(x) \leqslant p_l$, $x \in \bar{X}^*$, cannot be deleted whatever small r_m is used in the deletion operator $f(x_i^m) - p_l > r_m$. Obviously, $\bar{X}^\circ \subset \bar{X}^*$ since $p^\circ < p_l$. Take a number α such that $p^\circ < \alpha < p_l$ and define the set $X_\alpha = \{x \mid f(x) < \alpha,\ x \in \bar{X}\}$ which is non-empty for every $\alpha \in (p^\circ, p_l)$ since $\bar{X}^\circ \subset X_\alpha \subset \bar{X}^* \subseteq \bar{X}$ and $\bar{X}^\circ \neq \emptyset$. Since f is continuous and $\bar{X}$ is robust, the set X_α is, by the lemma, also robust and, thus, its interior int X_α is non-empty. Because $N \geqslant 2$ and partitions are made for every m, $m \to \infty$, and successive subcubes $\bar{C}_i^m$ are becoming smaller with the edge $c_m = c/N^m \to 0$ as $m \to \infty$, for a sufficiently big $m = m(\alpha)$ there will exist a subcube $\bar{C}_{i_0}^{m(\alpha)}$ completely contained in the int X_α. For the grid point $x_{m(\alpha)}$ of this subcube we have $x_{m(\alpha)} \in X_\alpha$ so that by construction of X_α necessarily $f(x_{m(\alpha)}) < \alpha < p_l$, contradicting the non-appearance of grid points $x_m \in \bar{X}$ with $f(x_m) = p_m < p_l$. Thus, descent in (4.8) does not cease although temporary stopovers may occur. This completes the proof of (4.10).

(d) *Existence and Nature of the Limit in* R^n

We have proved in (b) that $\bar{X}^\circ \subseteq \bigcap_{m=0}^{\infty} \bar{B}_m$. It remains to prove the inverse inclusion. Denote $\bar{V} = \bigcap_{m=0}^{\infty} \bar{B}_m$, which is closed and non-empty since $\bar{X}^\circ \subseteq \bar{V}$ and $\bar{X}^\circ \neq \emptyset$. Let $\bar{W} = \bar{V} \cap \bar{X}$. Obviously, $\bar{X}^\circ \subseteq \bar{W}$, so $\bar{W} \neq \emptyset$. Take any point $\tilde{x} \in \bar{W}$. Since $\tilde{x} \in \bar{W} = (\bigcap_{m=0}^{\infty} \bar{B}_m) \cap \bar{X}$, so $\tilde{x} \in \bar{B}_m$ for all $m = 1, 2, \ldots$, whence $f(\tilde{x}) - p_{m-1} \leqslant r_m$, $\forall m$, and, thus, $p^\circ = \min_{x \in \bar{X}} f(x) \leqslant f(\tilde{x}) \leqslant p_{m-1} + r_m \to p^\circ$ since $r_m \to 0$, $p_m \to p^\circ$ as $m \to \infty$. We obtain that $f(\tilde{x}) = p^\circ$ for all $\tilde{x} \in \bar{W}$, which means that $\bar{W} \subseteq \bar{X}^\circ$. Due to the inclusion $\bar{X}^\circ \subseteq \bar{W}$ proved before, we have $\bar{W} = \bar{X}^\circ$.

Further, in the procedure of the Beta-Algorithm we have after the first exclusion by the distinction operator (see (4.1), (4.2)) that $\bar{S}_0 = \bar{Y}_0$, where $\bar{Y}_0$ is the set containing $\bar{X}$ plus the uncertainty band $\mathscr{D}_0$ in the sense of Section 3. Similarly, in the second and further iterations we have $\bar{S}_m = \bar{Y}_m$, $m = 1, 2, \ldots$, whereby certain subcubes may be discarded from time to time by the deletion operator (4.3) based on the cost function and working independently.

By construction and due to intermittent action of the two operators, we have

$$\bar{C} = \bar{B}_0 \supseteq \bar{S}_0 \supseteq \bar{B}_1 \supseteq \bar{S}_1 \supseteq \bar{B}_2 \supseteq \bar{S}_2 \supseteq \cdots \supseteq \bar{B}_m \supseteq \bar{S}_m \supseteq \cdots. \tag{4.16}$$

If we denote $\bar{X}_m = \bar{B}_{m+1} \cap \bar{X}$, then by (4.16) we have

$$\bar{Y}_m \supseteq \bar{X}_m \quad \text{and} \quad \bar{X}_{m+1} \subseteq \bar{X}_m, \; m = 0, 1, \ldots . \tag{4.17}$$

Moreover, due to the distributive property of the intersection and to the facts $\bar{X}^\circ \neq \varnothing$, $\bar{W} = (\bigcap_{m=0}^{\infty} \bar{B}_m) \cap \bar{X} = \bar{X}^\circ$, we obtain that

$$\bigcap_{m=0}^{\infty} \bar{X}_m = \bar{X}^\circ \neq \varnothing. \tag{4.18}$$

Since the edges of subsequent cubes $c_m = c/N^m \to 0$, we are in the setting of Definition 3 (see (3.7) to (3.11)). The fact of intermittent exclusions by another operator (4.3) does not cause any harm. Indeed, each subcube $\bar{C}_i^1 \in \bar{S}_0$ discarded by (4.3) consists of two portions

$$\bar{C}_i^1 = P_i^d + \bar{P}_i^x, \tag{4.19}$$

where $\bar{P}_i^x = \bar{C}_i^1 \cap \bar{X}$ and $P_i^d = \bar{C}_i^1 - \bar{P}_i^x$. The portion $\bar{P}_i^x$, if non-empty, is accounted for in the set-valued sequence $\bar{X}_m$, which according to Definition 3 may vary arbitrarily under two conditions: $\bar{X}_{m+1} \subseteq X_m$ and $\bigcap_{m=0}^{\infty} X_m \neq \varnothing$, which are satisfied in our case (see (4.17), (4.18)). If the portion P_i^d is non-empty, then

$$P_i^d \subset \bar{S}_0 - \bar{B}_1 \cap \bar{X} = \bar{Y}_0 - \bar{X}_0 = \mathscr{D}_0 \tag{4.20}$$

and its exclusion by deletion operator (4.3) only helps the action of the distinction operator. This situation holds for all iterations $m = 0, 1, 2, \ldots$.

Now, by (4.16) we have

$$\bigcap_{m=0}^{\infty} \bar{Y}_m = \bigcap_{m=0}^{\infty} \bar{S}_m = \bigcap_{m=0}^{\infty} \bar{B}_m = \bar{V}, \tag{4.21}$$

so the limit $\lim_{m \to \infty} \bar{Y}_m = \bigcap_{m=0}^{\infty} \bar{Y}_m$ is defined.

The distinction operator applied in (4.1) is *precise*, which according to (3.9), (3.11), (4.18), (4.21), means

$$\bar{V} = \bigcap_{m=0}^{\infty} \bar{Y}_m = \lim_{m=0} \bar{Y}_m = \lim_{m=0} \bar{X}_m = \bigcap_{m=0}^{\infty} \bar{X}_m = \bar{X}^\circ. \tag{4.22}$$

Thus, $\bar{V} = \bar{W} = \bar{X}^\circ$, which completes the proof. ∎

Remark. The operations (4.1)–(4.2), (4.3)–(4.5), and (4.6)–(4.7) are interchangeable and the structure of the algorithm can be adjusted to the specifics of a particular problem. For example, the speed of convergence may increase, if (4.6)–(4.7) is done first and then p_1 of (4.7) is used instead of p_0 in (4.3). If the particular realization of the semi-certain distinction operator is complex, then it is advantageous to put (4.1)–(4.2) at the end of Iteration 1.

Consider now the case when the distinction operator is imprecise or unavailable at all, the latter meaning that there are no exclusions by $\varphi_i = 0$, so that $J_0 = I_0$ in (4.1), $\bar{S}_0 = \bar{B}_0$ in (4.2), and $\bar{S}_m = \bar{B}_m$, $m = 1, 2,...,$ in (4.16).

THEOREM 2. *If the Beta-Algorithm includes an imprecise distinction operator or no distinction operator, then*

$$\lim_{m \to \infty} p_m = p^\circ = \min_{x \in \bar{X}} f(x), \tag{4.23}$$

$$\lim_{m \to \infty} x_m = x^\circ \in \bar{X}^\circ = \{x \mid f(x) = p^\circ, x \in \bar{X}\}, \tag{4.24}$$

$$\bar{X}^\circ \subseteq \bigcap_{m=0}^{\infty} \bar{B}_m. \tag{4.25}$$

Proof. The statement (4.23), identical to (4.10) of Theorem 1, was proved in (a), (b), and (c) above without usage of the precision of the distinction operator, so it is correct. The statement (4.24) follows from the construction of the marginal comparison constant generator (4.7), from the ultimately non-ceasing descent in (4.8) and from the compactness of $\bar{X}$. The statement (4.25) follows from the non-elimination of global minimizers proved in (b). ∎

The Beta-Algorithm without a distinction operator will be called the Reduced Beta-Algorithm. This algorithm is of slower convergence and does not generally determine the set $\bar{X}^\circ$ of all global minimizers but only some larger set $\bar{V} = \bigcap_{m=0}^{\infty} \bar{B}_m$ containing $\bar{X}^\circ$.

5. CONCLUSIONS

We see that the Beta-Algorithm always finds the global minimum p° of $f(x)$ over a compact robust set $\bar{X}$ and at least one of its global minimizers no matter whether it is equipped with a distinction operator of whatever quality. If it is equipped with a *precise* distinction operator, then it also finds the minimizing set $\bar{X}^\circ = \{x \mid f(x) = p^\circ, \ x \in \bar{X}\}$. If the distinction operator is imprecise, then the algorithm delivers some larger set $\bar{V}$ containing $\bar{X}^\circ$.

In the case of a quasi-cubic set $\bar{X}$ the Beta-Algorithm is not needed since the simpler cubic algorithm [1] will do the job. For the cubic algorithm distinction operators are redundant and the marginal comparison constant generator (4.7) coincides with the extremal generator (1.12) proposed in [1].

Acknowledgment

The paper was written at the Chateau du Mont-Sainte-Anne, Beaupré, Québec, during the student ski-week. It is instructive that the idea of the semi-certain distinction operator and the principal lines of the algorithm came to the author in the chairlift on the way to the summit, January 7–8, 1986.

References

1. E. A. Galperin, The cubic algorithm, *J. Math. Anal. Appl.* **112**, No. 2 (1985), 635–640.
2. E. A. Galperin, Two alternatives for the cubic algorithm, *J. Math. Anal. Appl.* **125**, No. 1 (1987), 229–237.